

## Population Genetics

- One obstacle to understanding evolution is the common misconception that organisms evolve, in a Darwinian sense, in their lifetimes.
  - Natural selection *does* act on individuals by impacting their chances of survival and their reproductive success.
  - But the evolutionary impact of natural selection is *only apparent in tracking how a population of organisms changes over time.*
- It is the population, not its individual, that evolve.
- Evolution on the scale of populations, called *microevolution*, is defined as a change in the allele frequencies in a population.
- *The Origin of the Species* convinced most biologists that species are the products of evolution, but acceptance of natural selection as the main mechanism of natural selection was more difficult
  - What was missing in Darwin's explanation was an understanding of inheritance that could explain how chance variations arise in a population while also accounting for the precise transmission of these variations from parents to offspring...enter Gregor Mendel...
- When Mendel's research was rediscovered in the early twentieth century, many geneticists believed that the laws of inheritance conflicted with Darwin's theory of natural selection.
  - Darwin emphasized quantitative characters, those that vary along a continuum.
  - These characters are influenced by multiple loci.
  - Mendel and later geneticists investigated discrete "either-or" traits.
- An important turning point for evolutionary theory was the birth of **population genetics**, which emphasizes the extensive genetic variation within populations and recognizes the importance of quantitative characters.
  - Advances in population genetics in the 1930s allowed the perspectives of Mendelism and Darwinism to be reconciled.
    - This provided a genetic basis for variation and natural selection.
- A **population** is a localized group of individuals that belong to the same species.
  - One definition of a **species** (among others) is a group of populations whose individuals have the potential to interbreed and produce fertile offspring in nature.
- The total aggregate of genes in a population at any one time is called the population's **gene pool**.
  - It consists of all alleles at all gene loci in all individuals of a population.
  - Each locus is represented twice in the genome of a diploid individual.
    - Individuals can be homozygous or heterozygous for these homologous loci.
  - If all members of a population are homozygous for the same allele, that allele is said to be *fixed*.
  - Often, there are two or more alleles for a gene, each contributing a relative frequency in the gene pool.

## A population's gene pool is defined by its allele frequencies

- Imagine a wildflower population with two flower colors.
  - The allele for red flower color ( $R$ ) is completely dominant to the allele for white flowers ( $r$ ).
- Suppose that in an imaginary population of 500 plants, 20 have white flowers (homozygous recessive -  $rr$ ).
  - The other 480 plants have red flowers.
    - Some are heterozygotes ( $Rr$ ), others are homozygous ( $RR$ ).
    - Suppose that 320 are  $RR$  and 160 are  $Rr$ .
- Because these plants are diploid, in our population of 500 plants there are 1,000 copies of the gene for flower color.
  - The dominant allele ( $R$ ) accounts for 800 copies ( $320 \times 2$  for  $RR$  +  $160 \times 1$  for  $Rr$ ).
  - The frequency of the  $R$  allele in the gene pool of this population is  $800/1000 = 0.8$ , or 80%.
  - The  $r$  allele must have a frequency of  $1 - 0.8 = 0.2$ , or 20%.

## The Hardy-Weinberg Theorem describes a non evolving population

- The **Hardy-Weinberg theorem** describes the gene pool of a *non evolving* population.
- This theorem states that the frequencies of alleles and genotypes in a population's gene pool will remain constant over generations unless acted upon by agents *other than* Mendelian segregation and recombination of alleles.
  - *The shuffling of alleles after meiosis and random fertilization should have no effect on the overall gene pool of a population.*
- In our imaginary wildflower population of 500 plants, 80% (0.8) of the flower color alleles are  $R$  and 20% (0.2) are  $r$ .
- How will meiosis and sexual reproduction affect the frequencies of the two alleles in the next generation?
  - We assume that fertilization is completely random and all male-female mating combinations are equally likely.
- Because each gamete has only one allele for flower color, we expect that a gamete drawn from the gene pool at random has a 0.8 chance of bearing an  $R$  allele and a 0.2 chance of bearing an  $r$  allele.
- Using the rule of multiplication, we can determine the frequencies of the three possible genotypes in the next generation.
  - For the  $RR$  genotype, the probability of picking two  $R$  alleles is 0.64 ( $0.8 \times 0.8 = 0.64$  or 64%).
  - For the  $rr$  genotype, the probability of picking two  $r$  alleles is 0.04 ( $0.2 \times 0.2 = 0.04$  or 4%).
  - Heterozygous individuals are either  $Rr$  or  $rR$ , depending on whether the  $R$  allele arrived via sperm or egg.
    - The probability of ending up with both alleles is 0.32 ( $0.8 \times 0.2 = 0.16$  for  $Rr$ ,  $0.2 \times 0.8 = 0.16$  for  $rR$ , and  $0.16 + 0.16 = 0.32$  or 32% for  $Rr + rR$ ).
- As you can see, the processes of meiosis and random fertilization have maintained the same allele and genotype frequencies that existed in the previous generation.

- For the flower-color locus, the population's genetic structure is in a state of equilibrium, **Hardy-Weinberg equilibrium**.
  - Theoretically, the allele frequencies should remain at 0.8 for R and 0.2 for r forever.
- The Hardy-Weinberg theorem states that the processes involved in a Mendelian system have no tendency to alter allele frequencies from one generation to another.
  - The repeated shuffling of a population's gene pool over generations cannot increase the frequency of one allele over another.
- The Hardy-Weinberg theorem also applies to situations in which there are three or more alleles and with other interactions among alleles than complete dominance.
- Generalizing the Hardy-Weinberg theorem, population geneticists use ***p*** to represent the frequency of one allele and ***q*** to represent the frequency of the other allele.
  - The combined frequencies must add to 100%; therefore  $p + q = 1$ .
  - If  $p + q = 1$ , then  $p = 1 - q$  and  $q = 1 - p$ .
- In the wildflower example *p* is the frequency of red alleles (*R*) and *q* of white alleles (*r*).
  - The probability of generating an *RR* offspring is  $p^2$  (an application of the rule of multiplication).
    - In our example,  $p = 0.8$  and  $p^2 = 0.64$ .
  - The probability of generating an *rr* offspring is  $q^2$ .
    - In our example,  $q = 0.2$  and  $q^2 = 0.04$ .
  - The probability of generating *Rr* offspring is  $2pq$ .
    - In our example,  $2 \times 0.8 \times 0.2 = 0.32$ .
- The genotype frequencies should add to 1:
 
$$p^2 + 2pq + q^2 = 1$$
  - For the wildflowers,  $0.64 + 0.32 + 0.04 = 1$ .
- This general formula is the **Hardy-Weinberg equation**.
- Using this formula, we can calculate frequencies of alleles in a gene pool if we know the frequency of genotypes or the frequency of genotypes if we know the frequencies of alleles.
- We can use the Hardy-Weinberg theorem to estimate the percentage of the human population that carries the allele for a particular inherited disease, phenylketonuria (PKU) in this case.
  - About 1 in 10,000 babies born in the United States is born with PKU, which results in mental retardation and other problems if left untreated.
  - The disease is caused by a recessive allele.
- From the epidemiological data, we know that frequency of homozygous recessive individuals ( $q^2$  in the Hardy-Weinberg theorem) = 1 in 10,000 or 0.0001.
 

***\*\*always start with  $q^2$ , then all other values can be determined!***

  - The frequency of the recessive allele (*q*) is the square root of  $0.0001 = 0.01$ .
  - The frequency of the dominant allele (*p*) is  $p = 1 - q$  or  $1 - 0.01 = 0.99$ .
  - The frequency of carriers (heterozygous individuals) is  $2pq = 2 \times 0.99 \times 0.01 = 0.0198$  or about 2%.
- Thus, about 2% of the U.S. population carries the PKU allele.

- Populations at Hardy-Weinberg equilibrium must satisfy five conditions.
  - (1) *Very large population size*. In small populations, chance fluctuations in the gene pool, genetic drift, can cause genotype frequencies to change over time.
  - (2) *No migrations*. Gene flow, the transfer of alleles due to the movement of individuals or gametes into or out of our target population can change the proportions of alleles.
  - (3) *No net mutations*. If one allele can mutate into another, the gene pool will be altered.
  - (4) *Random mating*. If individuals pick mates with certain genotypes, then the mixing of gametes will not be random and the Hardy-Weinberg equilibrium does not occur.
  - (5) *No natural selection*. If there is differential survival or mating success among genotypes, then the frequencies of alleles in the next variation will deviate from the frequencies predicted by the Hardy-Weinberg equation.
- Evolution usually results when any of these five conditions are not met - when a population experiences deviations from the stability predicted by the Hardy-Weinberg theory.